# Genetic variation in RIN3 in the Belgian population supports its involvement in the pathogenesis of Paget's disease of bone and modifies the age of onset

# Genetic variation in *RIN3* in the Belgian population supports its involvement in the pathogenesis of Paget's disease of bone and modifies the age of onset.

*Raphaël De Ridder[1], Eveline Boudin[1], Geert Vandeweyer[1], Jean-Pierre Devogelaer[2], Erik Fransen[1],*

*Geert Mortier[1], Wim Van Hul,*[*1]*

[1] Center of Medical Genetics, University of Antwerp & Antwerp University Hospital, Antwerp, Belgium.
[2] Department of Rheumatology, Saint-Luc University Hospital, Université Catholique de Louvain, Brussels, Belgium

*Corresponding author:
prof. dr. Wim Van Hul
wim.vanhul @uantwerpen.be
Prins Boudewijnlaan 43, 2650 Edegem, Belgium

Declaration of interest: none to declare

## Abstract

Paget's disease of bone (PDB) is a common, late-onset bone disorder characterized by focal increase of bone turnover. Mutations in the *SQSTM1* gene are found in up to 40% of patients and recent GWAS have led to novel associations with several loci. *RIN3*, the candidate gene located at the associated 14q32 locus, has recently been studied in a British cohort to elucidate its contribution to the pathogenesis. In this study we performed a genetic screening of *RIN3* in an unrelated cohort to validate these findings and to further explore genetic variation in this gene in the context of PDB. In our screening, we examined the 5' untranslated region (UTR), the exonic regions and the intron-exon boundaries of the gene in a control cohort and a patient cohort. Our findings show clustering of variation similar to the British cohort and support a protective role for common genetic variation (rs117068593, p.R279C) in the proline-rich region and a functionally relevant role for rare genetic variation in the domains that mediate binding and activation of its interaction partner, Rab5. Additive regression models, fitted for the common variants, validated the association of the rs117068593 variant with the disease ($OR^{+/+}$:0.315; $OR^{+/-}$:0.562). In addition, our analyses revealed a potentially modifying effect of this variant on the age of onset of the disease. In conclusion, our findings support the involvement of genetic variation in *RIN3* in PDB and suggest a role for *RIN3* as a potential modifier of the age of onset of the disease.

## Keywords

Paget's disease of bone; pathogenesis; targeted sequencing; molecular inversion probes; *RIN3*; modifier

## 1. Introduction

Paget's disease of bone (PDB) is the second most prevalent bone disorder and is characterized by late onset development of focal lesions affecting one or several bones of the skeleton. These characteristic lesions arise due to osteoclastic hyperactivity, giving an initial osteolytic appearance on imaging. This resorptive phase is then followed by accelerated but disorganized bone formation. The aberrantly formed tissue is structurally inferior and susceptible to fracture and deformity, and lies at the basis of a wide range of complications. The presence of a positive family history in 5-40% of cases and a strong geographical component to its prevalence suggests an important role for genetic factors in the pathogenesis of PDB [1]. This hypothesis was confirmed with the identification of domain-specific mutations in the *sequestosome 1* (*SQSTM1*) gene [2,3]. These mutations are found in up to 40% of patients with a clear familial background, indicating that a significant fraction of the heritability remains unexplained [1]. In an effort to untangle the underlying genetic architecture genome wide association studies (GWAS) have been performed, resulting in the association of several new loci with the disease [4,5]. *RIN3*, found at the associated 14q32 locus, encodes for the Ras and Rab Interactor 3 (RIN3) protein. The protein has a highly similar organization compared to other members of the RIN family of proteins. Starting from the N-terminus, the Src Homology 2 (SH2) domain, followed by a Proline-rich region (PRR), the RIN homology (RH) domain, the vacuolar protein sorting 9 (VPS9) domain and finally the Ub-like Ras association (RA) domain can be distinguished (Figure 1) [6]. RIN3 has been identified as a guanine nucleotide exchange factor (GEF) and catalyzes the transition from the inactive GDP-bound state to the active GTP-bound conformation for a number of small GTPases (Rab5, Rabex, and Rab31) of the Rab5 subfamily [6,7]. These proteins play a central role in the regulation of intracellular membrane trafficking between distinct organelles by organizing a membrane domain through the recruitment of the effector proteins required for initiation of the transport process [8]. Rab5 coordinates the budding of clathrin-coated vesicles at the plasma membrane, facilitates transport of these vesicles to the early endosomes and enables homotypic fusion of these endosomes [9,10]. It is therefore considered a key regulator of endocytosis [11]. *In vitro* experiments provided preliminary evidence for a regulatory role for RIN3 in the early steps of this endocytic process and the specific functions carried out by RIN3's distinct domains. Furthermore, these studies have demonstrated that deletions and missense variants within the abovementioned domains of RIN3 and other GEFs have the potential to affect nucleotide exchange for the small GTPases and cytoplasmic localization upon activation [6,7,12]. Hence, genetic variation in *RIN3* may alter the regulation of the endocytic and transcytotic pathways in the osteoclast.

The precise function of RIN3 in bone biology is – so far- unknown, but its role as a potential modulator of the osteoclastic intracellular trafficking pathway makes *RIN3* an interesting subject for further studies. Recently, Vallet *et al.* undertook a first step towards unraveling the role of this gene in PDB pathobiology [13]. In their study, they identified a risk haplotype marked by the rs117068593C allele and a potentially functional variant (p.Y793H) in the VPS9 domain of the protein, which is quintessential in conferring RIN3's GEF activity. Based on these arguments, they propose that *RIN3* is the candidate gene for PDB at the 14q32 locus [13]. In our work, we aimed to perform a replication study in an independent population to confirm their findings and looked to further explore genetic variation in the *RIN3* gene that potentially contributes to the Pagetic phenotype.

## 2. Material and methods

### 2.1. Study population

Our genetic analyses were performed on DNA extracted from whole blood from 190 patients and 165 controls negative for causative mutations in the *SQSTM1* gene. Patients were diagnosed were based on total alkaline phosphatase levels and radiological or scintigraphic examinations. The control cohort consists of individuals that have no history of bone disease or non-traumatic fractures. All individuals included in this study are of Belgian ancestry and were screened for *SQSTM1*-mutations simultaneous to our screening for genetic variation in the *RIN3* gene to exclude false negatives. One patient was found positive for the p.P392L mutation in SQSTM1 and was excluded from further analyses. The basic characteristics of our control and patient cohorts can be found in table 1.

### 2.2. Targeted sequencing and Sanger validation.

Targeted enrichment of *RIN3* was performed using the molecular inversion probe (MIP) technology. Using the previously described MIPGEN pipeline, we designed probes for the 5' untranslated region (5' UTR), coding region and the splice regions of *RIN3* [14]. After MIP capture, enrichment, and indexing for the individual samples based on the protocol previously described by O'Roak *et al.* [15-17], the samples were pooled. The pools were diluted to 1.7pM prior to sequencing on the Illumina NextSeq 500 Next Generation Sequencing platform using the 2 x 75bp Mid-output flow cell (Illumina, San Diego, California, USA).

As a positive control, our gene panel also included MIPs for enrichment of the *SQSTM1* gene. The VCF files were generated using an in-house available pipeline. In summary, data analysis consisted of mapping the read-pairs to the human reference genome (hg19) using BWA v0.7.4, after which overlapping fragments within the read-pair were trimmed. Variants were called using multi-sample variant calling with the Unified Genotyper tool (GATK

v3.5.0). Using the resulting VCF files, variant filtering and annotation was carried out using VariantDB [18]. A prediction of the functional effect of the variants was made using several *in silico* prediction tools, including the Combined Annotation Dependent Depletion (CADD) tool v1.3 and REVEL (rare exome variant ensemble learner). The CADD scoring system results in a Phred-like scaled prediction score, which estimates the deleteriousness compared to all genetic variation. A CADD-score of 20 indicates that the variant is predicted to be in the top 1% most deleterious variants, a CADD-score of 30 representing the top 0.1% [19]. The REVEL tool uses a combination of 13 prediction programs to estimate the pathogenicity of a variant. The REVEL predictions yield a score between 0 and 1 and the higher scoring variants are more likely to be pathogenic [20].

The resulting variants were also confirmed using PCR amplification followed by enzymatic clean-up and Sanger sequencing. The primers for PCR amplification were designed using Primer3 (http://bioinfo.ut.ee/primer3/) based on the template sequence (accession code: NM_024832) The amplicons were amplified using the GoTaq® G2 DNA Polymerase (Promega Corporation, Madison, WI) and amplification of the PCR product was verified by agarose gel electrophoresis. Removal of unincorporated dNTPs and primers was performed by incubation in the presence of Exonuclease I (New England Biolabs, Inc., Ipswich, Massachusetts, USA) and Calf Intestinal Alkaline Phosphatase (CIAP, Roche Applied Science, Hoffmann-La Roche AG, Basel, Switzerland). Sanger sequencing was performed using the ABI Prism BigDye terminator cycle sequencing ready reaction kit (Applied Biosystems, Foster City, California, USA) using the same primers used for PCR amplification. Using the BigDye XTerminator purification kit we removed unincorporated BigDye terminators, prior to sequence determination using the ABI 310 Genetic Analyzer (Applied Biosystems, Foster City, California, USA). Upon coverage analysis, the first 1000bp of exon 6 of *RIN3* weren't covered properly in our NGS approach. Therefore, this region was also sequenced using the traditional Sanger sequencing-based workflow.

## 2.3. Statistical Analyses

Statistical analyses reported in this study have been performed in the computing environment R version 3.4.2.. Logistic regression models have been fitted to test for a significant association of alternate alleles with either disease state or the extent of disease as defined by the presence of Pagetic lesions in one (monostotic) or multiple (polyostotic) bones of the patients' skeletons. To test for an association between the genotype and the number of affected bones, quasi-Poisson regression models were fitted. Linear regression models were fitted to analyze potential effects of alternative allele dosage on patients' age of onset. The SNP genotype was entered as independent variable and sampling age was entered as covariate. To test the significance of the association between the genotype and the different outcome variables, a likelihood ratio test was carried out, comparing the model

described above to a null model only containing the sampling age. A linkage disequilibrium (LD)-plot (data not shown) was generated based on the genotyping data of our sequencing effort in European control and PDB cohorts using the HaploView software (Broad Institute, MA, USA). We observed substantial LD ($r^2$=0.79) between two SNPs that were commonly observed in our cohorts. To obtain a valid p-value that accounts for multiple testing of 6 variants in the presence of LD, we generated the empirical null distribution of the most significant p-value among the 6 SNPs tested here. We ran 10,000 Monte Carlo simulations, in each of which the outcome variable (disease status and age of onset respectively) was randomly permuted. Subsequently, the phenotype was regressed on the genotype for each separate SNP (using the same regression model as the true data), and the most significant of the 6 p-values was saved. The 10,000 p-values collected through these Monte Carlo simulations represent the distribution of the most significant p-value under the null hypothesis. We calculated the fraction of the p-values generated under the null hypothesis that was equal to or smaller than the observed minimal p-value obtained in the association tests between the real (unpermuted) phenotype and the 6 SNPs. This fraction represents the empirical p-value, which indicates a significant association accounting for multiple testing if this p-value is lower than 0.05

.

## 3. Results

When screening our cohorts for genetic variation occurring in the UTR, the coding sequence and the splice boundaries of RIN3, we identified 22 distinct variants. Of these, 12 were missense variants, 7 were synonymous variants, and 3 variants were located in the UTR's. Based on the combined results of the prediction tools used for our variants, we identified 6 rare variants that were predicted to be deleterious and that were found exclusively in one of the two cohorts (Table 2).

Among these, 2 rare variants (p.P16L and p.W63C) are located in the N-terminal region of RIN3 and are found in 3 control individuals (Figure 1). Of special interest is the p.W63C (rs150221413) variant in the SH2 domain, which is predicted to be deleterious by all prediction tools that were used (Table 2). The integrative prediction tool CADD ranks this variant as belonging to the top 0.43% most damaging variants with a CADD score of 23.60. A second integrative prediction tool, REVEL, which is focused on missense variation also marks this variant as a potentially functional variant with a score of 0.52. A second series of rare variants (p.K689R, p.Y793H, p.K838T and p.R859C) were exclusively found in a total of 5 patients. Contrary to the N-terminal variants in the control cohort, these four variants reside in the C-terminal RH and VPS9 domains that govern the binding of the Rab5 GTPase and catalysis of the GTP nucleotide exchange, respectively (Figure 1). CADD ranks these as belonging to the top

0,18%-0.03% most deleterious variants (Table 2). Elevated scores predicted by REVEL supports their potential for a functional effect.

Similar to the findings reported by Vallet *et al.*, we identified a cluster of common variation occurring in the PRR of RIN3 [13]. Interestingly, one of these common variants (rs117068593; p.R279C), is also predicted to be deleterious by most prediction tools used in our study indicating that the variant may have a functional effect [13]. Also, in the additive regression models that we fitted for the 6 common variants that we identified, this variant was the only common variant to show statistically significant differences (Table 3). The allele frequencies differed significantly (p=0.010) between patients and our control population, and within our patients we observed an association (p=0.014) of the alternative allele with the age of disease onset. This association between allele frequency and disease status remained significant after an empirical correction for multiple testing accounting for presence of LD, using 10,000 Monte-Carlo simulations (p=0.048). The association between allele frequency and age of disease onset, which was nominally significant without correction, still showed a trend towards significance after the empirical multiple testing correction (p=0.067) (Table 3.).

## 4. Discussion

Recent GWAS into the genetic architecture of Paget's disease of bone resulted in the identification of a number of new candidate genes, including *RIN3* [4,5]. To date the precise function of RIN3 in bone biology and homeostasis is unknown, but its association with lower and upper limb bone mineral density (BMD) suggests a significant role for RIN3 in the regulation of bone mass development and maintenance of skeletal health [21]. This hypothesis is supported by the involvement of RIN3 as a GEF in modulating the activity of members of the Rab family of small GTPases. Longstanding evidence shows the importance of these GTPases in the intracellular transport pathway [11], and their involvement in regulating osteoclast maturation and activity [22,23]. However, a more profound understanding of the function of many Rabs and their GEFs is still lacking, especially in the osteoclast. Similar to *SQSTM1*, variation in *RIN3* has also been reported to be associated with neurodegenerative disease [24,25]. The common basis in these associations can be found in the regulation of autophagic processes, in which targeting of vesicles between two cellular compartments by Rab proteins plays an important role [26,27]. In 2015, a study by Vallet *et al.* was the first to explore the contribution of genetic variation in the *RIN3* gene in the pathogenesis of PDB [13]. Our study was aimed at validating their findings and further expanding the repertoire of variation in *RIN3* in the context of PDB, resulting in the identification of 22 variants (Figure 1). Of these variants, 3 were found outside of the coding sequence in the UTRs. As described earlier, we identified several regions in the protein in

which potentially interesting variation occurs. It is to be noted that our study is limited to *in silico* predictions, further functional evaluation of these variants would be of interest.

In the N-terminal SH2 domain, we identified the p.W63C substitution in 2 control individuals. These individuals, and the control population at large, have no history of skeletal disease or non-traumatic fracture. The substitution involves a highly conserved nucleotide, resulting in the change of an aromatic tryptophan residue to a cysteine residue which is scored as the most evolutionary distant by the Grantham method [28]. Previous studies demonstrated that loss of function in the SH2 domain has clear effects on GEF activity of RIN3 [7]. This supports the idea that a deleterious variant in this domain could affect the endocytic pathway. Secondly, SH2 domains are protein interaction sites that typically bind phosphotyrosine-containing proteins, including phosphorylated tyrosine receptor kinases as is the case for RIN1 [29]. Altered interaction patterns could also affect the function of RIN3. Altered transport dynamics in the osteoclast could have a protective effect in the two control individuals who carried the variant. As a potentially functional variant in the RIN3 SH2 domain was described in one of the patients in the British population [13], further research would be of interest to explore this duality. Both effects on RIN3 protein structure or changes in protein-protein interactions with the SH2 domain could be implicated.

A second series of rare variants (p.K689R, p.Y793H, p.K838T and p.R859C) were exclusively found in patients. There is no indication that the individuals carrying these variants have a more severe phenotype, with a limited amount of bones affected and an average age of onset, except for the patients carrying the p.K838T and p.R859C variants that have an age of onset of 57 and 48 years, which are lower than the average age of onset observed in our cohort. Based on protein structure and homology models described in literature and SWISS-MODEL protein structure homology-modelling [30], the p.K689R variant is predicted to be located in the vicinity of the end of the αHB4 helix of the helical bundle of the RH domain. Deletion of this homologous region in a splice variant of RIN1 prevents its binding with Rab5 [31,32]. The p.Y793H, p.K838T and p.R859C variants are predicted to be located in the αV4, αV6 and αC helices of the VPS9 domain. The structural characterization of this evolutionarily conserved VPS9 domain in Rabex-5, a second GEF for Rab5, demonstrated the importance of the αV4 and αV6 helices in the interaction with Rab5 and of the αC helix in soluble expression of the protein [31]. Vallet *et al*. suggest that one of the rare variants we also identified in our patient cohort, p.Y793H, affects the structural stability of the protein [13]. Altered catalytic activity, soluble expression, and structural stability could all contribute to disruption of the Rab-cycle.

Besides these rare variants, a third cluster of common variants was found in exon 6 that primarily affects the intrinsically disordered Pro-rich region of the protein (Figure 1) [33]. This observation shows marked similarity to the variant distribution in the work reported by Vallet *et al.*. Of specific interest in this cluster is the p.R279C (rs117068593) substitution. The alternate allele occurs in a significantly higher rate in the control cohort as compared to our PDB cohort, decreasing the risk for PDB ($OR^{+/+}$:0.315; $OR^{+/-}$:0.562), and suggesting a protective effect for this common variant. This protective effect is also evident within the patient cohort, where our statistical analyses indicate a potential modifying effect of this variant on the age of onset, increasing the age of onset by 5.336 years per alternative allele that the patient carries (Table 3). In the British cohort, this variant was highly significantly associated with disease state and through haplotype analysis Vallet *et al.* reported the presence of a risk haplotype (rs10498635C–rs117068593C) on which 96% of all rare variants occurred, and it was suggested that the risk allele acts as a marker for rare variants [13]. We successfully replicated the previously reported association and expanded on their findings by demonstrating a modifying effect on the age of onset of the phenotype. In addition to these associations, our *in silico* prediction tools also show similar results and suggest a potential functional effect of this variant [13]. The intrinsically disordered PRR is involved in interactions with other proteins (e.g. amphiphysin II) [6]. Transition of disorder to order in intrinsically disordered regions can affect protein functionality, but in the literature only a subtle shift towards increased order of the PRR has been reported for this variant [34,13]. Whether the p.R279C variant itself underlies the protective effect observed in both independent populations and the phenotype-modifying effect we observe in our patient cohort, or whether variation occurring on the haplotype contributes collectively towards the reported effects is currently unknown.

In conclusion, our findings support the involvement of genetic variation in *RIN3* in PDB and suggest a role for *RIN3* as a potential modifier of the age of onset of the disease.

## Acknowledgements

The authors would like to thank Alex Hoischen (Radboudumc Nijmegen) for his help in the probe design used in our targeted sequencing approach.

## Funding

This work was supported by grants of the ‘Fonds voor Wetenschappelijk Onderzoek Vlaanderen’ (FWO grant G019712N and G031915N) and the European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 602300 (SYBIL). RDR holds a doctoral grant with the Fonds voor Wetenschappelijk Onderzoek Vlaanderen (1S07717N). EB and GV hold postdoctoral grants (12A3814N and 12D1717N) with the Fonds voor Wetenschappelijk Onderzoek Vlaanderen.

## Compliance with Ethical Standards

Ethical approval and informed consent by all patients were obtained prior to the study. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

## References


1. Alonso N, Calero-Paniagua I, Del Pino-Montes J (2017) Clinical and Genetic Advances in Paget's Disease of Bone: a Review. Clinical reviews in bone and mineral metabolism 15 (1):37-48. doi:10.1007/s12018-016-9226-0
2. Hocking LJ, Lucas GJ, Daroszewska A, Mangion J, Olavesen M, Cundy T, Nicholson GC, Ward L, Bennett ST, Wuyts W, Van Hul W, Ralston SH (2002) Domain-specific mutations in sequestosome 1 (SQSTM1) cause familial and sporadic Paget's disease. Human molecular genetics 11 (22):2735-2739
3. Laurin N, Brown JP, Morissette J, Raymond V (2002) Recurrent mutation of the gene encoding sequestosome 1 (SQSTM1/p62) in Paget disease of bone. American journal of human genetics 70 (6):1582-1588. doi:10.1086/340731
4. Albagha OM, Wani SE, Visconti MR, Alonso N, Goodman K, Brandi ML, Cundy T, Chung PY, Dargie R, Devogelaer JP, Falchetti A, Fraser WD, Gennari L, Gianfrancesco F, Hooper MJ, Van Hul W, Isaia G, Nicholson GC, Nuti R, Papapoulos S, Montes Jdel P, Ratajczak T, Rea SL, Rendina D, Gonzalez-Sarmiento R, Di Stefano M, Ward LC, Walsh JP, Ralston SH, Genetic Determinants of Paget's Disease C (2011) Genome-wide association identifies three new susceptibility loci for Paget's disease of bone. Nature genetics 43 (7):685-689. doi:10.1038/ng.845
5. Albagha OME, Visconti MR, Alonso N, Langston AL, Cundy T, Dargie R, Dunlop MG, Fraser WD, Hooper MJ, Isaia G, Nicholson GC, Montes JD, Gonzalez-Sarmiento R, di Stefano M, Tenesa A, Walsh JP, Ralston SH (2010) Genome-wide association study identifies variants at CSF1, OPTN and TNFRSF11A as genetic risk factors for Paget's disease of bone. Nature genetics 42 (6):520-U526. doi:10.1038/ng.562
6. Kajiho H, Saito K, Tsujita K, Kontani K, Araki Y, Kurosu H, Katada T (2003) RIN3: a novel Rab5 GEF interacting with amphiphysin II involved in the early endocytic pathway. Journal of cell science 116 (Pt 20):4159-4168. doi:10.1242/jcs.00718
7. Kajiho H, Sakurai K, Minoda T, Yoshikawa M, Nakagawa S, Fukushima S, Kontani K, Katada T (2011) Characterization of RIN3 as a guanine nucleotide exchange factor for the Rab5 subfamily GTPase Rab31. The Journal of biological chemistry 286 (27):24364-24373. doi:10.1074/jbc.M110.172445
8. Stenmark H, Olkkonen VM (2001) The Rab GTPase family. Genome biology 2 (5):REVIEWS3007.1-3007.7. doi: 10.1186/gb-2001-2-5-reviews3007
9. Gorvel JP, Chavrier P, Zerial M, Gruenberg J (1991) rab5 controls early endosome fusion in vitro. Cell 64 (5):915-925
10. Bucci C, Wandingerness A, Lutcke A, Chiariello M, Bruni CB, Zerial M (1994) Rab5a Is a Common Component of the Apical and Basolateral Endocytic Machinery in Polarized Epithelial-Cells. P Natl Acad Sci USA 91 (11):5061-5065. doi:DOI 10.1073/pnas.91.11.5061
11. Bucci C, Parton RG, Mather IH, Stunnenberg H, Simons K, Hoflack B, Zerial M (1992) The Small Gtpase Rab5 Functions as a Regulatory Factor in the Early Endocytic Pathway. Cell 70 (5):715-728. doi:Doi 10.1016/0092-8674(92)90306-W
12. Yoshikawa M, Kajiho H, Sakurai K, Minoda T, Nakagawa S, Kontani K, Katada T (2008) Tyr-phosphorylation signals translocate RIN3, the small GTPase Rab5-GEF, to early endocytic vesicles.

Biochemical and biophysical research communications 372 (1):168-172. doi:10.1016/j.bbrc.2008.05.027
13. Vallet M, Soares DC, Wani S, Sophocleous A, Warner J, Salter DM, Ralston SH, Albagha OME (2015) Targeted sequencing of the Paget's disease associated 14q32 locus identifies several missense coding variants in RIN3 that predispose to Paget's disease of bone. Human molecular genetics 24 (11):3286-3295. doi:10.1093/hmg/ddv068
14. Boyle EA, O'Roak BJ, Martin BK, Kumar A, Shendure J (2014) MIPgen: optimized modeling and design of molecular inversion probes for targeted resequencing. Bioinformatics 30 (18):2670-2672. doi:10.1093/bioinformatics/btu353
15. O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, Carvill G, Kumar A, Lee C, Ankenman K, Munson J, Hiatt JB, Turner EH, Levy R, O'Day DR, Krumm N, Coe BP, Martin BK, Borenstein E, Nickerson DA, Mefford HC, Doherty D, Akey JM, Bernier R, Eichler EE, Shendure J (2012) Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science 338 (6114):1619-1622. doi:10.1126/science.1227764
16. Jansen S, Hoischen A, Coe BP, Carvill GL, Van Esch H, Bosch DGM, Andersen UA, Baker C, Bauters M, Bernier RA, van Bon BW, Claahsen-van der Grinten HL, Gecz J, Gilissen C, Grillo L, Hackett A, Kleefstra T, Koolen D, Kvarnung M, Larsen MJ, Marcelis C, McKenzie F, Monin ML, Nava C, Schuurs-Hoeijmakers JH, Pfundt R, Steehouwer M, Stevens SJC, Stumpel CT, Vansenne F, Vinci M, van de Vorst M, Vries P, Witherspoon K, Veltman JA, Brunner HG, Mefford HC, Romano C, Vissers L, Eichler EE, de Vries BBA (2018) A genotype-first approach identifies an intellectual disability-overweight syndrome caused by PHIP haploinsufficiency. European journal of human genetics : EJHG 26 (1):54-63. doi:10.1038/s41431-017-0039-5
17. Neveling K, Mensenkamp AR, Derks R, Kwint M, Ouchene H, Steehouwer M, van Lier B, Bosgoed E, Rikken A, Tychon M, Zafeiropoulou D, Castelein S, Hehir-Kwa J, Tjwan Thung D, Hofste T, Lelieveld SH, Bertens SM, Adan IB, Eijkelenboom A, Tops BB, Yntema H, Stokowy T, Knappskog PM, Hoberg-Vetti H, Steen VM, Boyle E, Martin B, Ligtenberg MJ, Shendure J, Nelen MR, Hoischen A (2017) BRCA Testing by Single-Molecule Molecular Inversion Probes. Clinical chemistry 63 (2):503-512. doi:10.1373/clinchem.2016.263897
18. Vandeweyer G, Van Laer L, Loeys B, Van den Bulcke T, Kooy RF (2014) VariantDB: a flexible annotation and filtering portal for next generation sequencing data. Genome medicine 6 (10):74. doi:10.1186/s13073-014-0074-6
19. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M (2018) CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Research:gky1016-gky1016. doi:10.1093/nar/gky1016
20. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, Cannon-Albright LA, Teerlink CC, Stanford JL, Isaacs WB, Xu J, Cooney KA, Lange EM, Schleutker J, Carpten JD, Powell IJ, Cussenot O, Cancel-Tassin G, Giles GG, MacInnis RJ, Maier C, Hsieh C-L, Wiklund F, Catalona WJ, Foulkes WD, Mandal D, Eeles RA, Kote-Jarai Z, Bustamante CD, Schaid DJ, Hastie T, Ostrander EA, Bailey-Wilson JE, Radivojac P, Thibodeau SN, Whittemore AS, Sieh W (2016) REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. American journal of human genetics 99 (4):877-885. doi:10.1016/j.ajhg.2016.08.016
21. Kemp JP, Medina-Gomez C, Estrada K, St Pourcain B, Heppe DH, Warrington NM, Oei L, Ring SM, Kruithof CJ, Timpson NJ, Wolber LE, Reppe S, Gautvik K, Grundberg E, Ge B, van der Eerden B, van de Peppel J, Hibbs MA, Ackert-Bicknell CL, Choi K, Koller DL, Econs MJ, Williams FM, Foroud T, Zillikens MC, Ohlsson C, Hofman A, Uitterlinden AG, Davey Smith G, Jaddoe VW, Tobias JH, Rivadeneira F, Evans DM (2014) Phenotypic dissection of bone mineral density reveals skeletal site specificity and facilitates the identification of novel loci in the genetic regulation of bone mass attainment. PLoS genetics 10 (6):e1004423. doi:10.1371/journal.pgen.1004423
22. Itzstein C, Coxon FP, Rogers MJ (2011) The regulation of osteoclast function and bone resorption by small GTPases. Small GTPases 2 (3):117-130. doi:10.4161/sgtp.2.3.16453
23. Weivoda MM, Oursler MJ (2014) The Roles of Small GTPases in Osteoclast Biology. Orthopedic & muscular system : current research 3. doi:10.4172/2161-0533.1000161

24. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thorton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Moron FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fievet N, Huentelman MW, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuiness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, St George-Hyslop P, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossu P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Deniz Naranjo MC, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, European Alzheimer's Disease I, Genetic, Environmental Risk in Alzheimer's D, Alzheimer's Disease Genetic C, Cohorts for H, Aging Research in Genomic E, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannefelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH, Jr., Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RF, Passmore P, Montine TJ, Bettens K, Rotter JI, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JS, Boerwinkle E, Riemenschneider M, Boada M, Hiltuenen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nothen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM, Van Broeckhoven C, Moskvina V, Seshadri S, Williams J, Schellenberg GD, Amouyel P (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nature genetics 45 (12):1452-1458. doi:10.1038/ng.2802

25. Fecto F, Yan J, Vemula SP, Liu E, Yang Y, Chen W, Zheng JG, Shi Y, Siddique N, Arrat H, Donkervoort S, Ajroud-Driss S, Sufit RL, Heller SL, Deng HX, Siddique T (2011) SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. Archives of neurology 68 (11):1440-1446. doi:10.1001/archneurol.2011.250

26. Jain N, Ganesh S (2016) Emerging nexus between RAB GTPases, autophagy and neurodegeneration. Autophagy 12 (5):900-904. doi:10.1080/15548627.2016.1147673

27. Xu W, Fang F, Ding J, Wu C (2018) Dysregulation of Rab5-mediated endocytic pathways in Alzheimer's disease. Traffic' was changed to 'Xu W, Fang F, Ding J, Wu C (2018) Dysregulation of Rab5-mediated endocytic pathways in Alzheimer's disease. Traffic. 19(4): 253–262. doi:10.1111/tra.12547

28. Grantham R (1974) Amino acid difference formula to help explain protein evolution. Science 185 (4154):862-864

29. Barbieri MA, Kong C, Chen PI, Horazdovsky BF, Stahl PD (2003) The Src homology 2 domain of Rin1 mediates its binding to the epidermal growth factor receptor and regulates receptor endocytosis. Journal of Biological Chemistry 278 (34):32027-32036. doi:10.1074/jbc.M304324200

30. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic acids research 46 (W1):W296-W303. doi:10.1093/nar/gky427

31. Delprato A, Merithew E, Lambright DG (2004) Structure, exchange determinants, and family-wide rab specificity of the tandem helical bundle and Vps9 domains of Rabex-5. Cell 118 (5):607-617. doi:10.1016/j.cell.2004.08.009

32. Tall GG, Barbieri MA, Stahl PD, Horazdovsky BF (2001) Ras-activated endocytosis is mediated by the Rab5 guanine nucleotide exchange activity of RIN1. Dev Cell 1 (1):73-82. doi:Doi 10.1016/S1534-5807(01)00008-9

33. Babu MM (2016) The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. Biochemical Society transactions 44 (5):1185-1200. doi:10.1042/BST20160172

34. Vacic V, Markwick PR, Oldfield CJ, Zhao X, Haynes C, Uversky VN, Iakoucheva LM (2012) Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder. PLoS computational biology 8 (10):e1002709. doi:10.1371/journal.pcbi.1002709
35. Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4 (7):1073-1082. doi:10.1038/nprot.2009.86
36. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S (2010) Identifying a High Fraction of the Human Genome to be under Selective Constraint Using GERP plus. PLoS computational biology 6 (12). doi:ARTN e1001025

10.1371/journal.pcbi.1001025
37. Adzhubei I, Jordan DM, Sunyaev SR (2013) Predicting functional effect of human missense mutations using PolyPhen-2. Current protocols in human genetics Chapter 7:Unit7 20. doi:10.1002/0471142905.hg0720s76
38. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, Deflaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG, Exome Aggregation C (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature 536 (7616):285-291. doi:10.1038/nature19057
39. Kajiho H, Fukushima S, Kontani K, Katada T (2012) RINL, guanine nucleotide exchange factor Rab5-subfamily, is involved in the EphA8-degradation pathway with odin. PLoS One 7 (1):e30575. doi:10.1371/journal.pone.0030575

## Tables

**Table 1 Characteristics of cohorts included for analysis**. The table details the number of individuals in our cohorts and their mean age and standard deviations. The mean age of onset and the proportion of polyostotic disease is shown for our patient cohort. N.A. Not applicable

| | | Controls | | Patients | |
|---|---|---|---|---|---|
| | | Male | Female | Male | Female |
| Age | Individuals | 86 | 79 | 95 | 94 |
| | | 68.64±5.338 | 70.70±5.369 | 72.84±11.22 | 74.20±9.152 |
| Age of onset | Individuals | N.A. | N.A. | 72 | 73 |
| | | | | 60.25±11.853 | 63.92±13.217 |
| Polyostotic | Individuals | N.A. | N.A. | 82 | 83 |
| | | | | 55 (67.1%) | 46 (55.4%) |

**Table 2 Variants observed in the *RIN3* gene in our Belgian control and patient cohorts.** Minor allele frequencies (MAF) were looked up in the gnomAD database (5/12/2017). Abbreviations: not applicable (N.A.); benign (B); tolerated (T); possibly deleterious (PD); deleterious (D) [28,35-38,19,20]

| NM_024832.3 | | | | | | | | | | Individuals | | | | MAF EU | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | Controls (N=165) | | Patients (N=189) | | Non- | MAF |
| Position | Identifier | Amino acid change | Nucleotide change | CADD score | Grantham | Polyphen | SIFT | GERP | REVEL | +/- | +/+ | +/- | +/+ | Finnish | All |
| chr14:92980189 | rs547747275 | 5' UTR | c.-88C>A | 8.75 | N.A. | N.A. | N.A. | -4.43 | N.A. | 2 | 0 | 1 | 0 | 0.67% | 0.53% |
| chr14:92980256 | rs368389701 | 5' UTR | c.-21C>A | 11.55 | N.A. | N.A. | N.A. | -1.56 | N.A. | 0 | 0 | 1 | 0 | 0.26% | 0.26% |
| chr14:93022098 | . | p.P16L | c.C47T | 21.50 | 98 | B (0.01) | D (0.01) | 3.24 | 0,05 | 1 | 0 | 0 | 0 | 0.004% | 0.002% |
| chr14:93022240 | rs150221413 | p.W63C | c.G189T | 23.60 | 215 | PD (1.00) | D (0) | 5.51 | 0,52 | 2 | 0 | 0 | 0 | 0.12% | 0.08% |
| chr14:93107590 | rs34101393 | p.L150L | c.C448T | 14.94 | N.A. | N.A. | N.A. | 4.01 | N.A. | 19 | 2 | 26 | 3 | 6.62% | 5.81% |
| chr14:93118038 | rs3829947 | p.H215R | c.A644G | 4.00 | 29 | B (0) | T (0.43) | -5.52 | 0,02 | 76 | 55 | 100 | 54 | 56.18% | 48.47% |
| chr14:93118198 | rs3814830 | p.A268A | c.C804T | 9.59 | N.A. | N.A. | N.A. | 3.4 | N.A. | 34 | 6 | 47 | 2 | 15.28% | 20.82% |
| chr14:93118229 | rs117068593 | p.R279C | c.C835T | 23.90 | 180 | PD (0.93) | T (0.07) | 3.65 | 0,12 | 44 | 7 | 36 | 2 | 17.83% | 13.15% |
| chr14:93118369 | rs770038852 | p.H325H | c.T1127C | 1.46 | N.A. | N.A. | N.A. | -6.31 | N.A. | 1 | 0 | 1 | 0 | 0.0009% | 0.0013% |
| chr14:93118668 | rs3742717 | p.T425M | c.C1274T | 5.73 | 81 | B (0.01) | T (0.27) | -0.538 | 0,07 | 39 | 7 | 49 | 3 | 17.56% | 23.54% |
| chr14:93118669 | rs3742716 | p.T425T | c.G1275A | 8.66 | N.A. | N.A. | N.A. | 0.923 | N.A. | 70 | 17 | 99 | 20 | 32.07% | 31.34% |
| chr14:93118674 | rs74074811 | p.R427Q | c.G1280A | 0.002 | 43 | B (0.01) | T (0.87) | -6.79 | 0,011 | 0 | 0 | 1 | 0 | 0.08% | 2.25% |
| chr14:93118790 | rs139248637 | p.I466L | c.A1396C | 1.79 | 5 | B (0.01) | T (0.74) | -3.16 | 0,045 | 1 | 0 | 1 | 0 | 0.37% | 0.36% |
| chr14:93119232 | rs12434929 | p.G613A | c.G1838C | 2.80 | 60 | B (0.00) | T (0.85) | 1.27 | 0,043 | 2 | 0 | 5 | 0 | 0.80% | 1.79% |
| chr14:93119407 | rs3818321 | p.S671S | c.C2013T | 11.61 | N.A. | N.A. | N.A. | -9.17 | N.A. | 4 | 0 | 3 | 1 | 1.70% | 7.39% |
| chr14:93125545 | . | p.K689R | c.A2066G | 27.80 | 26 | PD (0.99) | D (0) | 5.84 | 0,319 | 0 | 0 | 1 | 0 | 0% | 0.0004% |
| chr14:93142861 | rs147042536 | p.Y793H | c.T2377C | 27.30 | 83 | PD (0.99) | D (0) | 4.66 | 0,576 | 0 | 0 | 2 | 0 | 0.70% | 0.40% |
| chr14:93151377 | rs746397902 | p.K838T | c.A2513C | 27.70 | 78 | PD (0.99) | D (0) | 3.93 | 0,389 | 0 | 0 | 1 | 0 | 0.004% | 0.002% |
| chr14:93151439 | rs751226648 | p.R859C | c.C2575T | 35.00 | 180 | PD (1.00) | D (0) | 5.08 | 0,540 | 0 | 0 | 1 | 0 | 0% | 0.005% |
| chr14:93154537 | . | p.G966G | c.T2898C | 3.43 | N.A. | N.A. | N.A. | 0 | N.A. | 1 | 0 | 0 | 0 | 0 % | 0.0005% |
| chr14:93154540 | rs71461983 | p.G967G | c.C2901T | 16.90 | N.A. | N.A. | N.A. | 0 | N.A. | 4 | 0 | 5 | 0 | 1.89% | 1.65% |
| chr14:93154608 | rs769146691 | 3' UTR | c.*11G>A | 6.37 | N.A. | N.A. | N.A. | -2.56 | N.A. | 1 | 0 | 0 | 0 | 0.013% | 0.006% |

**Table 3 Results of regression analyses for variants commonly occurring in our control and patient cohorts.** Effect sizes, standard errors and p-values for our additive regression models are given. Nominally significant results are shown in bold. Associations withstanding the empirical multiple testing correction are marked by an asterisk. (MAF: minor allele frequency)

| Identifier | Amino acid change | MAF (%) | | Disease state | Polyostotic disease | Number of affected bones | Age of onset |
|---|---|---|---|---|---|---|---|
| | | Controls | Patients | | | | |
| rs34101393 | p.L150L | 6.97% | 8.47% | 0.20±0.28 (p=0.464) | 0.53±0.41 (p=0.198) | 0.25±0.19 (p=0.202) | 1.65±2.35 (p=0.483) |
| rs3829947 | p.H215R | 56.36% | 55.03% | -0.05±0.16 (p=0.755) | 0.16±0.26 (p=0.530) | 0.08±0.13 (p=0.541) | -2.00±1.62 (p=0.217) |
| rs3814830 | p.A268A | 13.94% | 13.49% | -0.06±0.22 (p=0.790) | 0.26±0.38 (p=0.503) | -0.05± 0.18 (p=0.762) | 2.74±2.25 (p=0.225) |
| rs117068593 | p.R279C | 17.58% | 10.58% | **-0.58±0.23 (p=0.010)*** | 0.20±0.38 (p=0.605) | -0.30± 0.19 (p=0.106) | **5.62±2.27 (p=0.014)** |
| rs3742717 | p.T425M | 16.06% | 14.55% | -0.13±0.21 (p=0.533) | 0.30±0.38 (p=0.431) | -0.07± 0.17 (p=0.671) | 3.46±2.26 (p=0.128) |
| rs3742716 | p.T425T | 31.52% | 36.77% | 0.23±0.17 (p=0.167) | -0.45±0.28 (p=0.108) | -0.07± 0.14 (p=0.602) | 1.29±1.74 (p=0.459) |

## Figure

**Fig.1 Representation of variation occurring in RIN3**. Protein domains are marked by rectangles, the smaller rectangle represents the Pro-rich region. Variation commonly occurring in the RIN3 protein are marked above the domain structure, rare variants are given below it. Variants predicted to be deleterious are shown in bold, the potentially disease-modifying rs117068593 variant is marked with an asterisk. Domain structure based on the work of Kajiho *et al.* [39].